

A DETAILED STATISTICAL ANALYSIS OF THE DARK MATTER MASS PROFILES OF GALAXY CLUSTERS

OLE HOST, STEEN H. HANSEN

Dark Cosmology Centre, Niels Bohr Institute, University of Copenhagen, Juliane Maries Vej 30, DK-2100 Copenhagen, Denmark

(Dated: July 6, 2009)

Draft version July 6, 2009

ABSTRACT

The distribution of dark matter in the halos that surround galaxies and galaxy clusters has been probed observationally, theoretically, and in numerical simulations. Yet there is still confusion about which of several suggested parameterized models is the better representation, and whether these models are truly universal. We use the observed temperature and density profiles of the intracluster medium of 11 relaxed galaxy clusters to investigate mass models for the dark matter halo using a thorough statistical analysis. We make careful comparisons between two- and three-parameter models, including the issue of a universal third parameter. We find that, of the two-parameter models, the NFW is still the best representation, but we also find moderate statistical evidence that a generalized three-parameter NFW model with a freely varying inner slope is preferred, despite penalizing against the extra degree of freedom. There is a strong indication that this inner slope needs to be determined for each cluster individually, some clusters have central cores and others have steep cusps. This implies that X-ray observations do not support the idea of a universal inner slope, but perhaps show a hint of a dependence with redshift or mass. The mass-concentration relation of our sample is in reasonable agreement with the latest predictions based on numerical simulations.

Subject headings: dark matter — galaxies: clusters: general — X-rays: galaxies: clusters

1. INTRODUCTION

The potential of gravitationally bound structures in the Universe, ranging in size from dwarf galaxies to galaxy clusters, is sourced by a composite mass distribution of dark matter, baryonic matter in gas form, and collapsed objects such as stars in galaxies and galaxies in clusters. The investigation of these mass distributions entails a number of questions: what is the shape of the distributions? Is it universal across ten magnitudes of mass and at all redshifts? Does it depend on cosmology or on the merger history of the individual halos? Since the dominant component of relaxed structures is dark matter, much focus has been aimed at dark matter-only halos.

There is little theoretical understanding of the distribution of matter in a dark matter halo. The main developments have been found through numerical simulations of the formation of structure in the universe within a given cosmological model. Advances have been achieved through the improvement of numerical codes as well as the increase of raw computing power on one hand and a more refined understanding of which questions that need to be answered on the other. Perhaps the most fundamental idea that has come out of the numerical approach is that relaxed halos are (nearly) universal in many respects, including the distribution of matter (Navarro et al. 1997; Taylor & Navarro 2001) and the dynamical structure (Bullock et al. 2001a; Hansen & Moore 2006). However, the simulations have not been able to reach agreement about the exact behavior of the profiles in the innermost regions, where the limited force resolution of simulations sets a lower limit to the radial range that can be probed. Various authors claim that the logarithmic slope of the density profile reaches a value between -1 and -1.5 , perhaps dependent on mass or merger history, and there is also discussion whether the inner slope

is actually universal or not (Moore et al. 1998; Klypin et al. 2001; Navarro et al. 2004; Fukushige et al. 2004; Merritt et al. 2006; Graham et al. 2006; Gao et al. 2008). A further complication arises when the simulations are compared with observations since the gravitational potential of the baryonic component, which is very time consuming to model in the simulations, cannot be neglected in the center. This complication can in principle both change the slope of the dark matter profile as well as alter the total mass profile (Blumenthal et al. 1986; El-Zant et al. 2001; Gnedin et al. 2004; Sommer-Larsen & Limousin 2009).

Theoretical efforts are hampered by the fact that, even under the strongest simplifying assumptions, there are not enough constraints to obtain unique solutions to the collisionless Boltzmann equation (Binney & Tremaine 1987) which governs a dark matter structure. Instead, one can take phenomenological input from numerical simulations such as the density profile itself, the pseudo-phase space density (Taylor & Navarro 2001; Dehnen & McLaughlin 2005), or the density slope-velocity anisotropy relation (from which Hansen & Stadel (2006) predict an inner slope of 0.8), and implement this into a Jeans equation analysis to predict the consequences of the ‘inspired guess’ (see also Zait et al. (2008) and references therein). Alternatively one can attempt to model the formation history of the halo including major mergers and steady accretion (e.g., Ryden & Gunn (1987); Ascasibar et al. (2004); Salvador-Solé et al. (2007); Del Popolo (2009), and references therein). While these approaches typically yield results in rough agreement with simulations, the modeling can also explore the physical connection between the static and dynamic properties of the halo as well as offer constrained extrapolations which are not accessible in simulations.

Observationally, there is a strong discrepancy between the numerical results and the inferred mass distributions

in dwarf and low surface brightness galaxies, which are much shallower than predicted, the so-called cusp/core-problem (see, e.g., Salucci et al. (2003); Spekkens et al. (2005); Gilmore et al. (2007)). At the opposite end of the mass spectrum, galaxy clusters are typically found to be in rough agreement with the cuspy numerical simulations, but with even greater scatter for the inferred inner slope. There is also significant discussion about the type of model and number of parameters that are necessary in order to obtain an acceptable description of the data. One common method is based on mass modeling through weak or strong gravitational lensing, which can yield results which are in good agreement with numerical simulations (Broadhurst et al. 2005; Comerford et al. 2006; Limousin et al. 2008; Richard et al. 2009), but also profiles that are significantly shallower (Sand et al. 2004, 2008). Another method is based on X-ray observations of the intracluster medium (ICM) which is supported against gravitational collapse by its own pressure. Again, authors find a range of inner slopes (Ettori et al. 2002; Lewis et al. 2003; Zappacosta et al. 2006; Saha & Read 2009). For both lensing and X-ray studies most authors focus on only one or a few clusters, which of course makes it more difficult to assess the universality of the profiles on an observational foundation.

In the present work we take a sample of 11 highly relaxed clusters and use the measurements of the X-ray emitting gas to infer model-independent mass profiles. We then compare with a number of different models that have been applied as mass profiles in the literature, focusing on three key questions: Which parameterized model is the most successful? How many free parameters are needed to describe the data adequately? Is there evidence for a universal inner slope/shape-type parameter? We answer these questions using a detailed statistical analysis based on Bayesian inference where we use the Bayesian evidence (or marginal likelihood) to make judgments about which model is preferred by the data.

2. DENSITY PROFILE MODELS

Most models that are used for modeling the mass distribution in halos have been proposed or introduced as fitting formulae applied to the halos found in the numerical simulations. Hence these models are not theoretically well-founded but rather form a basis on which the predictions of numerical simulations can be compared with observations. Almost all of these models have two free parameters which determine the mass scale and the spatial extent of the halo, and these two parameters are specific to each halo. Some models have one or more additional parameters which determines the shape of the profile, and which may or may not be universal. Here we consider a number of two- and three-parameter models.

A whole class of models are ‘double power-laws’ which asymptote to power laws at very small and very large radii. These models can conveniently be summarized in Hernquist’s (α, β, γ) parametrization (Hernquist 1990; Zhao 1996),

$$\rho(r) = \rho_0 \left(\frac{r}{r_s} \right)^{-\alpha} \left[1 + \left(\frac{r}{r_s} \right)^\gamma \right]^{-\frac{\beta-\alpha}{\gamma}}, \quad (1)$$

where ρ_0 and r_s are scaling constants to be determined for each halo individually. The inner power-law slope

is α and the outer slope is β , while the width of the transition region is controlled by γ . We consider four such two-parameter profiles: the NFW (Navarro et al. 1997), the Dehnen-McLaughlin (Dehnen & McLaughlin 2005), the Hernquist, and the Moore profile (Moore et al. 1998). The properties of these models are summarized in Table 1.

We also consider three three-parameter models: Two are simply generalized NFW profiles where, in the first case, we allow the inner slope α to vary. The motivation for this slopeNFW model was already apparent from the introduction. The second case, transNFW, is also a generalization of the NFW where now the transition parameter γ is free. Such a profile can mimic a steeper inner slope by pushing the inner power law behavior closer to the center. The third profile is the Sérsic (or Einasto) profile (Sérsic 1963; Einasto 1969),

$$\rho(r) = \rho_s \exp \left(-2n \left[\left(\frac{r}{r_s} \right)^{1/n} - 1 \right] \right), \quad (2)$$

where the parameter n determines the shape of the profile. For $n = 4$ the de Vaucouleurs’ law describing the surface brightness of elliptical galaxies is recovered. The shape parameter is sometimes given as $\alpha_s = n^{-1}$. Recently, the Sérsic profile has been claimed to provide a better fit than the NFW to Milky Way-sized haloes formed in numerical simulations, and, interestingly, with a shape parameter that varies significantly from halo to halo (Salvador-Solé et al. 2007; Navarro et al. 2008).

We map the scale radius r_s and scale densities ρ_s or ρ_0 of each model to the model-independent parameters r_{-2} and ρ_{-2} , which are the radius at which the slope of the density profile is -2 and the density at that radius, respectively. This mapping makes comparison of the models easier and enables identical priors to be used in the statistical analysis in all models.

3. DATA ANALYSIS

We revisit the sample of 11 highly relaxed, low redshift ($z < 0.1$) galaxy clusters observed with *XMM-Newton* which we already used in Host et al. (2009) to measure the dark matter velocity anisotropy profile for the first time (see also Hansen & Piffaretti (2007)). The members of this sample were selected to appear close to round on the sky and not have strong features in the temperature and density profiles. The spectral analysis and deprojection of the X-ray data was carried out in Kaas-stra et al. (2004) and Piffaretti et al. (2005). The deprojection method was non-parametric, i.e. without any parametric modeling of the radial temperature or density profiles. The outcome, and the starting point for the present analysis, was estimates of the ICM temperature T_i and electron number density $n_{e,i}$ with associated uncertainties in six or seven radial bins, for each of the clusters.

Assuming hydrostatic equilibrium, the ICM gas traces the gravitational potential according to (Cavaliere & Fusco-Femiano 1978)

$$\frac{k_B T}{\mu m_H} \left(\frac{d \ln n_e}{d \ln r} + \frac{d \ln T}{d \ln r} \right) = -\frac{GM_{\text{tot}}(r)}{r}, \quad (3)$$

where $\mu = 0.6$ is the mean molecular weight of the ICM. Almost all of the cluster mass resides in dark

TABLE 1
DENSITY PROFILE MODELS

Model	(α, β, γ)	r_{-2}/r_s	ρ_{-2}/ρ_0	$\mu(x = r/r_s)$
NFW	(1, 3, 1)	1	$\frac{1}{4}$	$\ln(1+x) - x/(1+x)$
D&M	$(\frac{7}{9}, \frac{31}{9}, \frac{4}{9})$	$\frac{121}{169}$	0.0338	$\frac{9}{20}(1+x^{4/9})^{-5}$
Hernquist	(1, 4, 1)	$\frac{1}{2}$	$\frac{16}{27}$	$x^2/[2(1+x)^2]$
Moore	$(\frac{3}{2}, 3, 1)$	$\frac{1}{2}$	$\frac{8}{3\sqrt{3}}$	$2 \sinh^{-1}(\sqrt{x}) - 2\sqrt{x/(1+x)}$
slopeNFW	$(\alpha, 3, 1)$	$2 - \alpha$	$(2 - \alpha)^{-\alpha}(3 - \alpha)^{\alpha-3}$...
transNFW	(1, 3, γ)	1	$\frac{1}{4}$...
Sérsic	...	1	1	$8^{-n}e^{2n}n^{1-3n}\gamma(3n, 2nx^{1/n})$

NOTE. — Properties of the density profiles that we consider, including the (α, β, γ) specification, the relations between (r_{-2}, ρ_{-2}) and (r_s, ρ_0) , and the shape $\mu(r)$ of the mass profile $M(r) = 4\pi r_s^3 \rho_0 \mu(r)$, if analytical. $\gamma(a, x)$ is the lower incomplete gamma function, $\gamma(a, x) = \int_0^x t^{a-1} e^{-t} dt$.

matter and the ICM, and therefore the dark matter mass distribution can be determined through $M_{\text{DM}}(r) = M_{\text{tot}}(r) - M_{\text{ICM}}(r)$. The ICM mass profile is given straight-forwardly by the density $\rho_{\text{ICM}} = \mu m_H n_e$.

We calculate $M_{\text{DM}}(r_i)$ of each radial bin through a Monte Carlo (MC) analysis in order to propagate uncertainties accurately. In detail, the prescription for each MC realization is as follows: In each bin i the best estimates of T_i and $n_{e,i}$ are added to random numbers drawn from Gaussian distributions representative of the uncertainties δT_i and $\delta n_{e,i}$. In order to apply Equation (3), we estimate the logarithmic derivative of, e.g., T at the bin-radius r_i by the slope of the unique parabola that passes through $(\ln r_{i-1}, \ln T_{i-1})$, $(\ln r_i, \ln T_i)$, and $(\ln r_{i+1}, \ln T_{i+1})$. In this way we can calculate the total mass interior to r_i for that data realization. We subtract the gas mass, estimated through a five-point Newton-Cotes integration formula applied to the same realization of the density data, and we arrive at the dark matter mass $M_{\text{DM},i}$. We impose a number of checks to determine if the derived data realization is physically sensible: the ICM temperature and density must be greater than zero in all bins, the total mass profile must be increasing with radius, and the dark matter mass profile and derived density profile must also be everywhere positive. If these conditions are not met the entire realization is discarded. This process is repeated until $N = 5000$ realizations have been accepted. From these the sample mean of $\ln M_i$ in each bin is determined, as well as the sample covariance matrix with elements

$$C_{ij} = \frac{1}{1-N} \sum_k^N (\ln M_{ik} - \langle \ln M_i \rangle) (\ln M_{jk} - \langle \ln M_j \rangle), \quad (4)$$

where N is the number of Monte Carlo realizations. Even though we sample the ICM temperature and density in each bin independently, the covariance matrix is not diagonal since the derivatives and physical consistency checks induce bin-to-bin correlations in the accepted sample. We use the mean and covariance of $\ln M$ rather than M since, by inspection, the former is closer to being Gaussian distributed.

4. STATISTICAL ANALYSIS

We take a Bayesian approach to the statistical analysis and the usual starting point is the likelihood function, which we calculate in the following manner.

It requires less manipulation of the data to calculate the mass profile from the observations than to calculate the density profile. Therefore we integrate the density profile analytically or numerically for each model to obtain the mass distribution and compare with the data in mass space, not density space. Further, as mentioned above, we have found in the MC analysis that the mass samplings in each bin are close to being log-normally distributed. Therefore we construct the likelihood $\mathcal{L}(M_i) = \exp(-\chi^2/2)$ from the χ^2 function,

$$\chi^2 = \sum_{i,j} (\ln M_i - \ln M(r_i)) C_{ij}^{-1} (\ln M_j - \ln M(r_j)), \quad (5)$$

where $M(r_i)$ is the model mass profile at the radial centre r_i of bin i , and $\ln M_i$ and C_{ij} are determined by the MC analysis.

The main goal is to decide which model is the better representation of the data. We do this by calculating the Bayesian evidence of each model, which is a quantitative measure of the agreement between model and data (Trotta 2008). First we calculate the likelihoods of each model on a grid in the parameter space $\theta = (\log r_{-2}, \log \rho_{-2})$. Next, we construct the posterior probability distribution by combining the likelihood function with a prior probability distribution $\pi(\theta)$ resembling our knowledge of $\log r_{-2}$ and $\log \rho_{-2}$ before taking the data into account. We discuss the choice of prior below. We then integrate the posterior to obtain the Bayesian evidence,

$$E = \int_{\text{all}} d\theta \pi(\theta) \mathcal{L}(\theta, \ln M_i), \quad (6)$$

which is essentially the weighted average of the likelihood over the prior volume. The evidence of a model, given the data and a prior, quantifies how well that model explains the data. It is important to stress that the comparison is made over all of the prior volume, not just at the best fitting set of parameters. When comparing models the Bayes factor $B_{12} = E_1/E_2$ shows how much more (or less) probable model 1 is than model 2, in light of the data. Traditionally, this is gauged on Jeffrey's scale where a Bayes factor of $\ln B_{12} < 1$ is labeled 'inconclusive' evidence for model 1 over model 2 while 'weak', 'moderate', and 'strong' evidence corresponds to $\ln B_{12}$ values < 2.5 , < 5 , and > 5 , respectively.

We choose priors which are constant in the logarithms of r_{-2} and ρ_{-2} . The flat logarithmic prior is the uninfor-

TABLE 2
BAYES FACTOR $\ln B$ FOR THE TWO-PARAMETER MODELS, RELATIVE
TO THE NFW PROFILE

Cluster	z	D&M	Hernq.	Moore
A262	0.015	-2.0	0.9	-3.0
NGC533	0.018	-1.7	1.2	-3.0
A496	0.032	-1.4	0.6	-1.2
2A0335+096	0.034	0.5	-0.1	13.2
A2052	0.036	1.9	-0.3	5.8
MKW9	0.040	0.5	-0.1	1.4
MKW3s	0.046	1.8	-0.3	6.2
A4059	0.047	1.5	-0.4	9.5
Sérsic 159-3	0.057	-0.5	1.6	2.7
A1795	0.064	2.5	-0.5	17.9
A1837	0.071	0.5	-0.2	1.2
Total	...	3.6	2.4	51

NOTE. — A positive value of $\ln B$ indicates that the NFW profile is preferred over the considered model. Note that this does not imply any bias towards the NFW as the Bayes factor of any two other models is just the difference between the respective Bayes factors given here.

mative prior for scaling parameters (Trotta 2008) since it reflects ignorance about the magnitude of the parameter. However we restrict the range of the priors, so that we end up with top-hat priors in $\log r_{-2}$ and $\log \rho_{-2}$. As a reference point we first assume a top-hat prior relative to the best estimate of r_{2500} as determined in the MC analysis. (The scale radius r_{2500} is defined as the radius within which the mean density is 2500 times the critical density of the universe.) The top-hat prior in $\log r_{-2}$ ranges from 1.5 magnitudes below r_{2500} to 0.5 above. The basic idea behind this prior is that the transition or ‘roll’ of a model should occur close to r_{2500} , as it does in haloes in numerical simulations, and also to prevent the model from behaving as a simple power-law by pushing the transition from the inner to the outer power law far away from the range of the data. We emphasize that this is still a conservative prior, as current simulations typically resolve 2–3 radial magnitudes with r_{-2} located about one order of magnitude below the virial radius (Bullock et al. 2001b). The prior in ρ_{-2} is also a top-hat in the logarithm and a range of 10^{-26} – $10^{-21} \text{ kg m}^{-3}$, which in practice means that the likelihood is vanishingly small at the boundaries of the prior.

4.1. Two-parameter model results

The result of the model comparison is summarized in Table 2, where the NFW model is compared against each of the other two-parameter models. A positive Bayes factor indicates that the NFW model is preferred. This does not imply any bias on the NFW since any two models can be compared by subtracting the Bayes factors we give for them from one another. We find that, individually, the clusters yield strong constraints only against the Moore model, while the evidences for or against the D&M and Hernquist models are either weak or inconclusive on Jeffrey’s scale. If instead we consider the cumulative Bayes factor summed over the full sample, the NFW is found to be the preferred model overall, i.e., as a universal two-parameter profile our sample favors the NFW model. The Hernquist profile and the D&M profile are weakly and moderately disfavored, respectively, with cumulative Bayes factors of 2.4 and 3.6 while the Moore profile is convincingly ruled out with a factor of 51. The weak constraint on the Hernquist profile is not surprising

as data extending out to the virial radius would likely be needed to properly distinguish this model from the NFW.

In Table 3 we present the effects of varying the priors. The evidence against the D&M profile increases to the level of strong when we limit the range of the prior in $\log r_{-2}$ to the smaller interval $(-0.75, 0.25)$, while the Bayes factor is reduced slightly on the larger range $(-3, 3)$. The evidence also becomes strong if we choose top-hat priors in (r_{-2}, ρ_{-2}) instead of the logarithmic priors. Finally, the D&M model is disfavored slightly more if we apply a ‘soft’ Gaussian prior in $\log r_{-2}$. The Bayes factor for the Hernquist model is robust under the same variations, while the Moore profile is very strongly ruled out in all cases. We conclude that our two-parameter model selection results are stable against variation amongst reasonable choices of priors, which means that the data are of sufficient quality to make robust conclusions.

A more interesting issue to consider than the priors is that the preference for the NFW profile over the Hernquist and D&M profiles is somewhat susceptible to ‘jack-knife’ resampling: if we recompute the cumulative Bayes factor eleven times systematically leaving a single cluster out each time, there are a few cases where the strength of the evidence is reduced to inconclusive but also cases where it is increased to strong (against the D&M). This is largely due to the fact that our data sample is somewhat inhomogeneous in terms of the relative statistical uncertainty on the mass profile. For example, a comparison of the error bars of MKW9 with those of A1795 or Sérsic 159-3 (see Figure 1) immediately shows that the former is much less constraining than the latter two. This means that our sample is a mixture of strongly and weakly constraining clusters and this is reflected in Figure 2 where the contributions from individual clusters clearly varies. There appears to be a trend that the clusters A262, NGC533, and A496, which are the lowest redshift and some of the least massive in our sample, stand out by preferring the D&M and the Moore profile. However, such trends are just as likely spurious selection effects caused by the relatively small sample but could be investigated with a larger sample. The D&M profile can easily be preferred by clusters that also prefer the Moore profile since, by extending the transition region, the D&M profile can push the inner asymptotic power law well inside the radial range of the data.

Finally we compare with a standard goodness-of-fit test: the minimum χ^2 values for the models support our more detailed analysis: for a total of 53 degrees of freedom we get minimum χ^2 ’s of 81 for the NFW, 93 for the D&M, 82 for the Hernquist, and 190 for the Moore profile. Major contributions to these χ^2 values come from the two clusters MKW3s with $\chi^2 = 14.2$ and A4059 with $\chi^2 = 13.2$ for the NFW model and similar or larger values for the other models. The corresponding p -values imply that the D&M χ^2 is about 20 times less likely to have occurred by chance (if the D&M model is correct) than the NFW model is (if the NFW model is correct). Compare this with the Bayesian odds that the NFW is ~ 40 times more probable than the D&M. Note that the actual best-fits are slightly smaller since we evaluate the χ^2 on a grid instead of minimizing it with a dedicated search. The χ^2 values show that, also in terms of goodness-of-fit, our sample is rather inhomogeneous. The rather poor

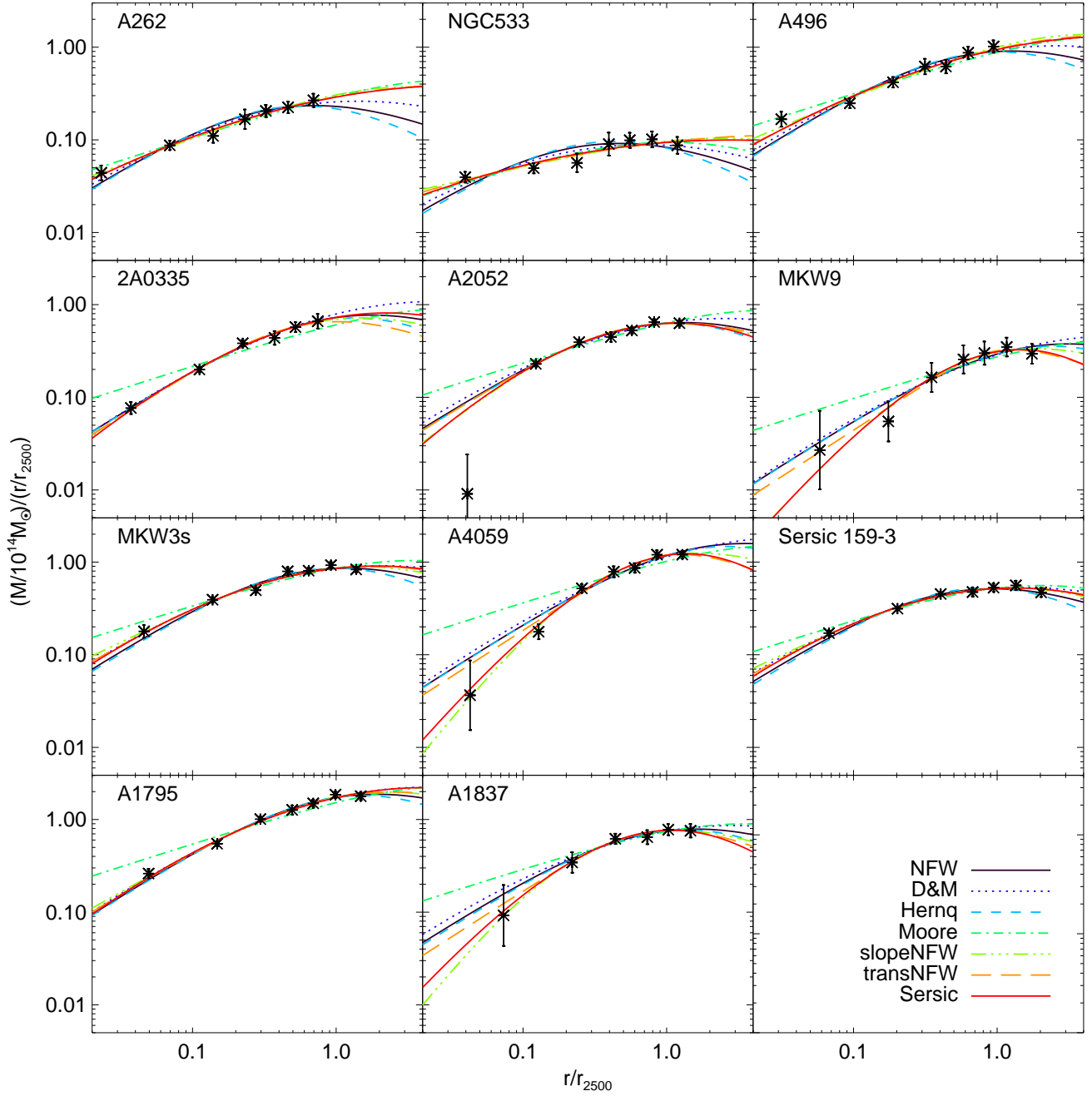


FIG. 1.— Mass profile of each cluster with 68% uncertainties and best-fit models. The radial axis has been scaled to the best estimate of r_{2500} from the MC analysis, and the mass axis has been scaled by r^{-1} so that the fitted models are approximately horizontal at r_{-2} .

total fit should not be judged too harshly since the halos in numerical simulations also show halo-to-halo scatter, which is not accounted for by the fitting profiles.

4.2. Three-parameter model results

For the three-parameter models we again want to test whether the models represent the data better than the NFW. In this case the comparison is slightly more involved since there is the freedom of an additional parameter to take into account. This naturally yields a lower value of the evidence if the extra parameter does not provide a better description of the data, or, more specifically, the third parameter must improve the fit over a

significant volume of parameter space in order to be preferred over the NFW. It is important to stress that there is no assumption about the third parameter being universal. On the contrary, we ask whether the data require the additional freedom of an extra parameter which is determined individually for each cluster.

The model comparison proceeds as before: we calculate the evidence for each of the three-parameter models with the same priors in $\log r_{-2}$ and $\log \rho_{-2}$ as in the fiducial two-parameter analysis for all models. For the slopeNFW we choose a top-hat prior for α which ranges from 0 to 1.75, i.e. from a cored profile to a profile slightly steeper than the Moore profile. We do not want to go all

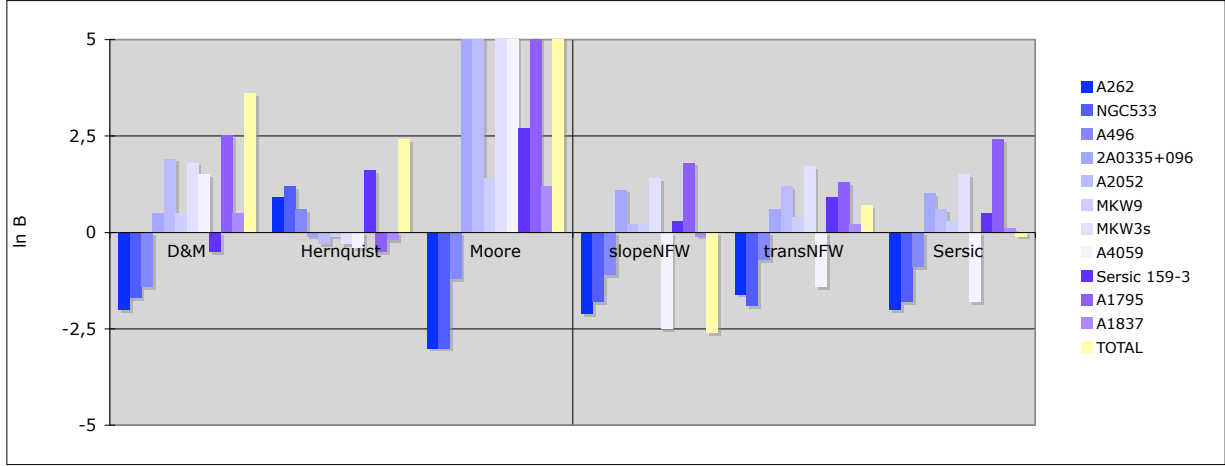


FIG. 2.— Bar chart of the Bayes factors $\ln B$ for the various models considered, relative to the NFW, as given in Table 2 and 4. The Bayes factors are additive so the contribution of individual clusters to the total Bayes factor is easily assessed. The values shown are based on the fiducial priors discussed in the text.

TABLE 3
TOTAL BAYES FACTOR $\ln B$ FOR THE TWO-PARAMETER MODELS
ASSUMING VARIOUS PRIORS, RELATIVE TO THE NFW PROFILE

Prior	Range $\log \rho_{-2}$	D&M	Hernq.	Moore
Top-hat in $\log r_{-2}$	(-1.5, 0.5)	3.6	2.4	51
Top-hat in $\log r_{-2}$	(-3, 3)	2.8	2.5	36
Top-hat in $\log r_{-2}$	(-0.75, 0.25)	6.8	2.3	63
Top-hat in (r_{-2}, ρ_{-2})	(-1.5, 0.5)	8.8	1.3	60
Gaussian in $\log r_{-2}$...	4.8	2.3	50

NOTE. — The top line is the fiducial prior used in Table 2. In the next two cases the range of the prior in $\log r_{-2}$ (in units of r_{2500} , see text) is varied, and in the following two cases a top-hat prior in r_{-2} and both r_{-2} and ρ_{-2} is applied. The final case assumes a Gaussian prior in $\log r_{-2}$ with mean -0.25 and width 0.5.

the way to -2 since r_{-2} tends to zero and eventually becomes undefined as α approaches -2 . For the transNFW, we choose a logarithmic prior with γ in the range (0.1, 4) which allows this profile to mimic a steeper inner profile by pushing the asymptotic inner power law inside the radial range of the data. Finally, we take a logarithmic prior for n in the range (2, 15) for the Sérsic model, motivated by numerical simulations which have best fits Sérsic profiles with $n=5-9$. The logarithmic prior has the advantage that it is invariant whether one prefers n or $\alpha_s = 1/n$ as the parameterization.

The resulting Bayes factors relative to the NFW are given in Table 4 and summarized in the chart in Figure 2. The individual clusters provide only weak evidence for or against any of the models. Based on the whole sample, the model selection is inconclusive for the transNFW and Sérsic models but there is ‘moderate’ evidence for the slopeNFW model over the NFW with a Bayes factor of -2.6 . This corresponds to odds of 13 to 1 in favor of the slopeNFW model and shows that, overall, the slopeNFW has the highest evidence E of all models considered. Hence the data show a moderate need for a free inner slope despite the penalty against the extra freedom built into the Bayesian analysis. It must be mentioned that most of the discriminatory power is carried by a few clusters such as NGC533, A4059, and A1795 and removal of any of these clusters from the sample would change the Bayes factor significantly. Therefore we

TABLE 4
BAYES FACTOR $\ln B$ FOR THE THREE-PARAMETER MODELS,
RELATIVE TO THE NFW PROFILE.

Cluster	slopeNFW	transNFW	Sérsic
A262	-2.1	-1.6	-2.0
NGC533	-1.8	-1.9	-1.8
A496	-1.1	-0.7	-0.9
2A0335+096	1.1	0.6	1.0
A2052	0.2	1.2	0.6
MKW9	0.2	0.4	0.3
MKW3s	1.4	1.7	1.5
A4059	-2.5	-1.4	-1.8
Sérsic 159-3	0.3	0.9	0.5
A1795	1.8	1.3	2.4
A1837	-0.1	0.2	0.1
Total	-2.6	0.7	-0.1

NOTE. — A positive value of $\ln B$ indicates that the NFW profile is preferred over the considered model. A top-hat prior in $\log r_{-2}$ of $(-1.5, 0.5)$ around the best estimate of r_{2500} for each cluster is assumed.

caution that the moderate preference for the slopeNFW model is somewhat susceptible to selection effects since, as noted above, the constraints from individual clusters vary in quality. We also find some sensitivity to the choice of prior: if the upper bound of α is extended from 1.75 up to 1.9, the Bayes factor for the slopeNFW model changes to -1.9 , while if it is set to the Moore profile at 1.5 the Bayes factor becomes -3.1 . If the lower bound of α is increased to 0.5, the Bayes factor remains relatively unchanged at -2.9 . While we believe that the fiducial priors used above are reasonable descriptions of the ‘state of knowledge’ based on numerical simulations, the sensitivity to the choice of prior indicates that the data do not necessarily confine the posterior to a sufficiently small region of the prior volume to provide unambiguous conclusions.

4.3. Constraints on the third parameters

Finally, for the three-parameter models we also want to constrain the preferred value of the third parameter. Unlike above, this analysis assumes that there is a universal value for the third parameter and attempts to identify that value. We use the same priors as in the previous analysis for each model, but now we marginal-

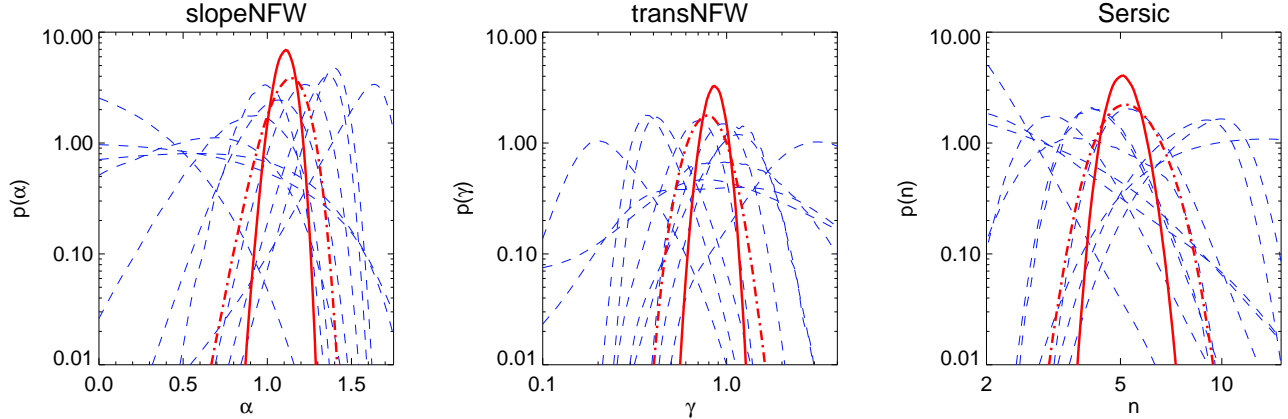


FIG. 3.— Probability distributions for the third parameter in each of the three-parameter models: slopeNFW α (left), transNFW γ (middle), and Sérsic n (right). In each panel, the full line shows the joint posterior for all clusters combined while the dot-dashed line shows the joint posterior obtained using the method of hyper-parameters (see text). The dashed lines are the pdf's of individual clusters. Note that each posterior is normalized to unity so it is not possible to draw conclusions about the quality of fit of the individual clusters from this plot. The standard 95% credible intervals are (1.00, 1.21) for α , (0.68, 1.06) for γ , and (4.3, 6.2) for n . With the hyper-parameters, the intervals are instead (0.85, 1.31) for α , (0.50, 1.28) for γ , and (3.5, 7.4) for n . We assume top-hat priors in α , $\ln \gamma$, and $\ln n$.

ize over the nuisance parameters ($\log r_{-2}$, $\log \rho_{-2}$) to find the one-dimensional posterior probability distribution for the third parameter for each cluster. Then we combine the results from the individual clusters into a joint posterior which is simply the product of the individual ones. We calculate 95% credible intervals for both the individual and the joint posterior. However, we know from the previous analysis that each three-parameter model is preferred by some clusters but not by others. Therefore we also use the method of hyper-parameters (Lahav et al. 2000) which allows the contribution from individual data-sets to the joint posterior to be weighted. These weights are marginalized over assuming logarithmic priors with the result that in the joint likelihood one replaces

$$\sum_i \chi_i^2 \rightarrow \sum_i N_i \ln \chi_i^2, \quad (7)$$

where N_i is the number of data points in data-set i . The upshot of all this is that clusters that are not described well by the model do not constrain the parameters as strongly as clusters that are well described. The price to pay is that the effective sample size is reduced which, all other things being equal, will lead to wider and more conservative credible intervals.

The results are shown in Figure 3, where in each panel the fully drawn line is the joint posterior, the dotted line is the hyper-parameters posterior, and the dashed lines are the posteriors of the individual clusters. The generalized NFW models are drawn slightly away from, but not in disagreement with, the NFW with 95% credible intervals of (1.00, 1.21) for α and (0.68, 1.06) for γ . The interval for the Sérsic n parameter is (4.3, 6.2), in good agreement with the values reported by the Aquarius numerical simulations for Milky Way-sized halos (Navarro et al. 2008). The intervals derived using the method of hyper-parameters are wider as expected: (0.85, 1.31) for α , (0.50, 1.28) for γ , and (3.5, 7.4) for n . The difference between the hyper-parameters method and the conventional calculation illustrates the need for a cautious approach to in-homogeneous data-sets. We believe the hyper-parameters yields the more trustworthy results in the case at hand, while on the other hand we acknowl-

edge that they are not very constraining.

An inspection of the contribution from individual clusters reveals some issues: It is clear that for each model a number of clusters provide very little information about the third parameter, i.e. the model describes the mass profile almost equally well regardless of the third parameter value. This is actually expected, given the varying size of the Bayes factors in table 4. There are also a few cases, particularly for the transNFW model, where the posterior peaks very close to or on the bounds of the prior. In such cases the results, e.g. the individual credible intervals, are of course very prior-dependent which again indicates that the data are not very discriminatory with respect to the prior. On the other hand, rather drastic priors or small sub-samples must be used in order to significantly affect the credible intervals of the joint posterior, especially for the hyper-parameters method.

Figure 4 shows the individual clusters' constraints on α , γ , and n . As could be expected given the varying nature of our results, there is perhaps the slightest of hints of a redshift-dependence in the constraints but our sample size does not allow us to probe such an issue in detail. It should again be noted that any hint of a redshift-dependence could actually be caused by a mass-dependence instead, since the two lowest redshift clusters in our sample are also the least massive.

A different picture emerges when we consider the overlap of the individual clusters' credible intervals for the slopeNFW model. For example, no value of α is contained in all 11 95% credible intervals, and only the very short range (1.08, 1.10) is contained in all but two intervals. Likewise the NFW $\alpha = 1$ case is excluded from four of the eleven intervals. These results, as well as a visual inspection of Figure 3, puts strong doubts about the concept of a universal shape parameter. The situation is not quite as compelling for the transNFW and Sérsic models which is likely the reason that they do not stand out from the NFW in the model selection. In fact, we believe it is a reasonable statement that the success of the slopeNFW model is precisely due to the very different preferred values of α from cluster to cluster. This puts a strong question mark against the idea of universal

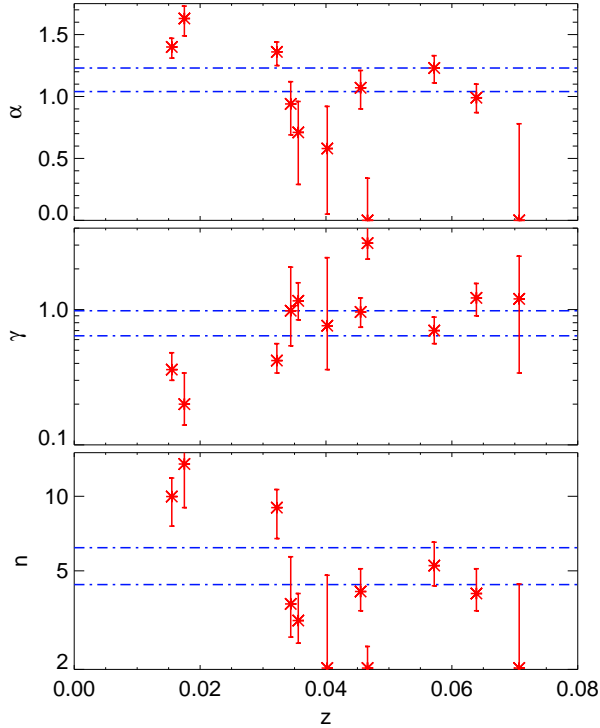


FIG. 4.— The individual clusters’ constraints on the third parameter in each of the three-parameter models. In this case we show the 68% credible intervals, and the horizontal lines indicate the 68% range of the joint posterior calculated using the method of hyper-parameters. Refer to Table 2 for the redshifts of each cluster.

third parameter.

We conclude that there is moderate evidence for the slopeNFW model to be preferred over the simple NFW, while the transNFW and Sérsic models do not stand out against the two-parameter NFW profile. If the inner slope of the slopeNFW model is universal, we constrain it to be close to -1 but preferably slightly steeper. However, the spread of the individual clusters’ preferred ranges suggests that the inner slope is not universal.

We also comment that the method of hyper-parameters method could in principle be extended to the model selection analysis. As a matter of fact, since the slopeNFW ‘contains’ both the NFW and the Moore models as subsets, we can derive the corresponding Bayes factor for the Moore profile which is only 10. This is of course a drastic reduction numerically but it does not alter the conclusion and anyway corresponds to rather convincing odds of about 20 000 : 1.

5. BIASES

So far we have discussed the interpretation of our results with respect to the statistical evidence. However, a number of biases, or systematic uncertainties, can be thought of that may affect our results. Loosely, these can be grouped into biases that affect both the individual cluster mass modeling and the combined analysis, and selection effects that only influence the latter.

The analysis rests on the ability to produce deprojected temperature and density profiles with uncertainties that correctly mirror the uncertainties in the spectral analysis of the X-ray data. This has been discussed extensively

in Kaastra et al. (2004). The basic assumption in determining the mass distribution of a galaxy cluster is that the cluster is relaxed, and obeys the equation of hydrostatic equilibrium. Numerical simulations indicate that the additional pressure associated with turbulence and bulk motion in the ICM yields an underestimate of the mass in the region of 5 – 20% with the larger values corresponding to large radii, r_{500} and greater (Nagai et al. 2007; Piffaretti & Valdarnini 2008; Lau et al. 2009). We do not expect this bias to exceed 10% in the present case since we do not model further out than to $\sim r_{2500}$. On the other hand, the same numerical simulations indicate that if the turbulent pressure is accounted for, an accurate mass reconstruction is possible. This point demonstrates that deviations from spherical symmetry are not a major concern in the error budget.

A related question is whether the parameterized profiles should be tested against the total mass distribution or the dark matter mass profile only. While the predictions of numerical simulations are founded in dark matter-only simulations, it is not clear how much a simulated dark matter-only mass profile is modified by the presence of baryons. Observationally, the ICM contributes about 5 – 15% of the total density in a cluster, again increasing with radius in the range of interest here, so formally there is a difference between the total and the dark matter profile’s radial dependence. To test the impact of this, we have rerun the statistical analyses described above without subtracting the ICM mass from the mass estimate of Equation (3). We find only minor differences: For the two-parameter models, the total Bayes factors relative to the NFW profile, assuming the fiducial prior as in Table 2, are 2.8 (D&M), 2.7 (Hernquist), and 54 (Moore), i.e. there is no significant change in the interpretation of the results. For the three-parameter models, the total Bayes factors become -3.1 (slopeNFW), 0.6 (transNFW), and -0.2 (Sérsic), which are in good agreement with the results in Table 4. Finally, the constraints on the third parameters for the three-parameter models are unchanged.

The fact that our results are stable whether we test the mass models against the total or dark matter-only mass profiles allows us to gauge how important the mass bias caused by turbulent pressure is. The point is that the turbulent pressure is expected to contribute the same amount (or less) to the total mass estimate as the ICM mass: both contributions are at the 5 – 15% level and radially increasing, and at the maximum radius we consider $\sim r_{2500}$ the gas fraction ($\sim 10\%$) is likely larger than the pressure bias. Since our results are the same whether we account for the ICM mass or not, we conclude that the systematic uncertainty is likely much smaller than the statistical uncertainty.

6. MASS-CONCENTRATION RELATION

An important consequence of the ‘bottom-up’ scenario of structure formation is that smaller halos are denser in the center, since they formed earlier when the density of the Universe was higher. This effect is seen in numerical simulations and it can be expressed as a relation between the halo mass and the concentration parameter. The concentration parameter is defined for a given overdensity as $c_{\Delta} = r_{\Delta}/r_{-2}$ (often r_s is used instead of r_{-2} but for the NFW this is unimportant). Simulations

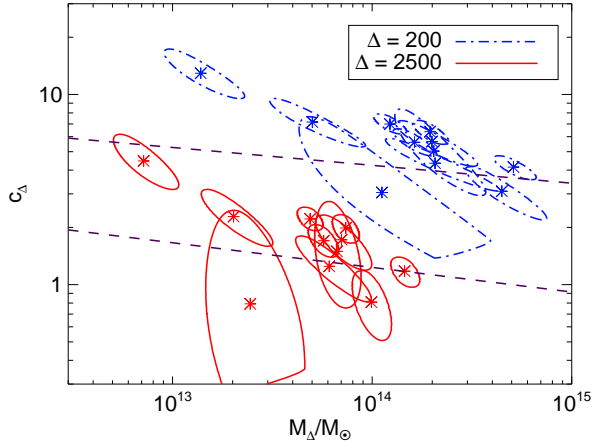


FIG. 5.— The mass-concentration relation of our sample, calculated within the NFW model. The contours contain 95% of the posterior PDF and are based on the fiducial prior. We show two contours for each cluster: (M_{2500}, c_{2500}) (red, full lines) which are derived within the radial range of the data, and (M_{200}, c_{200}) (blue, dot-dashed) which is based on an NFW model-dependent extrapolation to r_{200} . The dashed lines show the mean relations for the two values of Δ from the N-body simulations of Macciò et al. (2008), based on the WMAP5 cosmology. The relations are $\log c_{200} = 0.83 - 0.094 \log(M_{200}/10^{12} M_{\odot})$ and $\log c_{2500} = 0.35 - 0.130 \log(M_{2500}/10^{12} M_{\odot})$. Given the low redshift of our sample, we have not made any correction for a redshift evolution of c_{Δ} .

usually consider the mass-concentration relation at the virial radius r_{200} , but we can only reach that radius by model-dependent extrapolation. Therefore, in Figure 5, we show the mass-concentration relation of our sample calculated within the NFW model at both r_{2500} and extrapolated to r_{200} . We suggest that authors provide relations at both of these Δ 's as they complement each other in physical significance and observational accessibility.

As can be seen in Figure 5, our sample is not ideally suited to derive a relation from given that six sample members cluster at almost identical values of M_{Δ} . Instead we compare with the mass-concentration relation of the simulations presented in Macciò et al. (2008), which are in reasonable agreement with our sample except for the low mass NGC533. We emphasize that the orientation of the uncertainty ellipses is related only to the parameter degeneracies present in the combination of model and mass profile data and has nothing to do with the slope of the mass-concentration relation. The agreement between our observed mass-concentration relation and the predictions of numerical simulations resembles the recent X-ray analysis of Buote et al. (2007), but stands out from the significant discrepancy of the lensing study in Broadhurst et al. (2008).

7. SUMMARY & DISCUSSION

We have conducted a careful statistical analysis of the constraints on mass distribution models of galaxy clusters which can be derived from X-ray observations. We find that the NFW model is the preferred two-parameter model and that the Moore model is decisively ruled out. There is moderate evidence that the data require an additional free parameter that alters the shape of the mass profile, and the best choice is a model similar to the NFW but with a freely varying inner slope. If we assume this slope to be universal, we can constrain it to be close to or

slightly steeper than the NFW, but our data suggest that the shape parameter must be determined individually.

Significantly, the clusters in our sample prefer very different values for the inner slope, some prefer flat cores while others prefer steep cusps. The shape-parameters of the other two three-parameter models we consider, the Sérsic and transNFW, also show considerable scatter across our sample. We conclude that there is a strong indication in our data that the mass profile is not universal but suffers considerable halo-to-halo scatter. The limited size of our sample means that we cannot assess whether this is in disagreement with the results of numerical simulations. Of course, we can force universality of the inner slope, in which case we find that it is preferred to be slightly steeper than the NFW value of -1 . However, when the goodness-of-fit of each cluster is taken into account using the method of Bayesian hyper-parameters, the credible interval becomes significantly larger, partly because of the smaller effective sample size, but also because of the lack of universality.

This analysis stands out from the numerous observational results that claim significant discrepancies from simulations based on only one or a few observed clusters. We acknowledge that our sample size is still limited, but it allows us to discuss the issue of universality. Given that halos in numerical simulations which include baryons are still not readily mass produced with sufficient resolution, which makes the question of halo to halo scatter difficult to assess, it is not possible to decide if the strong indication of a non-universal model that we see is at odds with the numerical predictions.

Our results are largely insensitive to whether we compare the models with the dark matter mass profile or the total mass profile, and so we cannot judge whether one type of model is more appropriate for the dark matter halo or the total gravitational potential of a halo. There are two reasons for this: firstly, the uncertainties of primarily the ICM temperature profiles are too large, and secondly, the angular resolution in the center is not good enough. Of course the ICM is rather smoothly distributed and very good statistics would be needed for a model to fit either the total or dark matter distribution significantly better. For nine of the 11 clusters of our sample we readily found 2MASS (Skrutskie et al. 2006) cD or BCG galaxies very close to the X-ray center but including these in the mass budget does not make a difference, unless we assumed extreme mass-to-light ratios. This is again due to the limited angular resolution of the observations which implies that a large amount of dark matter is contained even within the radial center of the innermost bin.

The robustness of our results whether we use the dark matter or total mass profile gives us reasonable confidence that deviations from hydrostatic equilibrium in the ICM are not a major problem. Such a systematic uncertainty would yield a bias of at most 10–15% according to numerical simulations, which is similar to the difference between the total and dark matter mass profiles.

We thank Rocco Piffaretti for sharing data with us and for comments on the manuscript, and Andrea Macciò for providing the mass-concentration relation of his numerical simulations. The Dark Cosmology Centre is funded

by the Danish National Research Foundation.

Facilities: XMM

REFERENCES

- Ascasibar, Y., Yepes, G., Gottlöber, S., & Müller, V. 2004, *MNRAS*, 352, 1109
- Binney, J. & Tremaine, S. 1987, *Galactic Dynamics* (Princeton, NJ, Princeton University Press, 1987, 747 p.)
- Blumenthal, G. R., Faber, S. M., Flores, R., & Primack, J. R. 1986, *ApJ*, 301, 27
- Broadhurst, T., Takada, M., Umetsu, K., Kong, X., Arimoto, N., Chiba, M., & Futamase, T. 2005, *ApJ*, 619, L143
- Broadhurst, T., Umetsu, K., Medezinski, E., Oguri, M., & Rephaeli, Y. 2008, *ApJ*, 685, L9
- Bullock, J. S., Dekel, A., Kolatt, T. S., Kravtsov, A. V., Klypin, A. A., Porciani, C., & Primack, J. R. 2001a, *ApJ*, 555, 240
- Bullock, J. S., Kolatt, T. S., Sigad, Y., Somerville, R. S., Kravtsov, A. V., Klypin, A. A., Primack, J. R., & Dekel, A. 2001b, *MNRAS*, 321, 559
- Buote, D. A., Gastaldello, F., Humphrey, P. J., Zappacosta, L., Bullock, J. S., Brighenti, F., & Mathews, W. G. 2007, *ApJ*, 664, 123
- Cavaliere, A. & Fusco-Femiano, R. 1978, *A&A*, 70, 677
- Comerford, J. M., Meneghetti, M., Bartelmann, M., & Schirmer, M. 2006, *ApJ*, 642, 39
- Dehnen, W. & McLaughlin, D. E. 2005, *MNRAS*, 363, 1057
- Del Popolo, A. 2009, *ApJ*, 698, 2093
- Einasto, Y. I. 1969, *Astrophysics*, 5, 67
- El-Zant, A., Shlosman, I., & Hoffman, Y. 2001, *ApJ*, 560, 636
- Ettori, S., Fabian, A. C., Allen, S. W., & Johnstone, R. M. 2002, *MNRAS*, 331, 635
- Fukushige, T., Kawai, A., & Makino, J. 2004, *ApJ*, 606, 625
- Gao, L., Navarro, J. F., Cole, S., Frenk, C. S., White, S. D. M., Springel, V., Jenkins, A., & Neto, A. F. 2008, *MNRAS*, 387, 536
- Gilmore, G., Wilkinson, M. I., Wyse, R. F. G., Kleyna, J. T., Koch, A., Evans, N. W., & Grebel, E. K. 2007, *ApJ*, 663, 948
- Gnedin, O. Y., Kravtsov, A. V., Klypin, A. A., & Nagai, D. 2004, *ApJ*, 616, 16
- Graham, A. W., Merritt, D., Moore, B., Diemand, J., & Terzić, B. 2006, *AJ*, 132, 2701
- Hansen, S. H. & Moore, B. 2006, *New Astronomy*, 11, 333
- Hansen, S. H. & Piffaretti, R. 2007, *A&A*, 476, L37
- Hansen, S. H. & Stadel, J. 2006, *Journal of Cosmology and Astroparticle Physics*, 5, 14
- Hernquist, L. 1990, *ApJ*, 356, 359
- Host, O., Hansen, S. H., Piffaretti, R., Morandi, A., Ettori, S., Kay, S. T., & Valdarnini, R. 2009, *ApJ*, 690, 358
- Kaastra, J. S., Tamura, T., Peterson, J. R., Bleeker, J. A. M., Ferrigno, C., Kahn, S. M., Paerels, F. B. S., Piffaretti, R., Branduardi-Raymont, G., & Böhringer, H. 2004, *A&A*, 413, 415
- Klypin, A., Kravtsov, A. V., Bullock, J. S., & Primack, J. R. 2001, *ApJ*, 554, 903
- Lahav, O., Bridle, S. L., Hobson, M. P., Lasenby, A. N., & Sodrè, L. 2000, *MNRAS*, 315, L45
- Lau, E. T., Kravtsov, A. V., & Nagai, D. 2009, *arXiv:0903.4895*
- Lewis, A. D., Buote, D. A., & Stocke, J. T. 2003, *ApJ*, 586, 135
- Limousin, M., Richard, J., Kneib, J.-P., Brink, H., Pelló, R., Jullo, E., Tu, H., Sommer-Larsen, J., Egami, E., Michałowski, M. J., Cabanac, R., & Stark, D. P. 2008, *A&A*, 489, 23
- Macciò, A. V., Dutton, A. A., & van den Bosch, F. C. 2008, *MNRAS*, 391, 1940
- Merritt, D., Graham, A. W., Moore, B., Diemand, J., & Terzić, B. 2006, *AJ*, 132, 2685
- Moore, B., Governato, F., Quinn, T., Stadel, J., & Lake, G. 1998, *ApJ*, 499, L5+
- Nagai, D., Vikhlinin, A., & Kravtsov, A. V. 2007, *ApJ*, 655, 98
- Navarro, J. F., Frenk, C. S., & White, S. D. M. 1997, *ApJ*, 490, 493
- Navarro, J. F., Hayashi, E., Power, C., Jenkins, A. R., Frenk, C. S., White, S. D. M., Springel, V., Stadel, J., & Quinn, T. R. 2004, *MNRAS*, 349, 1039
- Navarro, J. F., Ludlow, A., Springel, V., Wang, J., Vogelsberger, M., White, S. D. M., Jenkins, A., Frenk, C. S., & Helmi, A. 2008, *arXiv:0810.1522*
- Piffaretti, R., Jetzer, P., Kaastra, J. S., & Tamura, T. 2005, *A&A*, 433, 101
- Piffaretti, R. & Valdarnini, R. 2008, *A&A*, 491, 71
- Richard, J., Pei, L., Limousin, M., Jullo, E., & Kneib, J. P. 2009, *A&A*, 498, 37
- Ryden, B. S. & Gunn, J. E. 1987, *ApJ*, 318, 15
- Saha, P. & Read, J. I. 2009, *ApJ*, 690, 154
- Salucci, P., Walter, F., & Borriello, A. 2003, *A&A*, 409, 53
- Salvador-Solé, E., Manrique, A., González-Casado, G., & Hansen, S. H. 2007, *ApJ*, 666, 181
- Sand, D. J., Treu, T., Ellis, R. S., Smith, G. P., & Kneib, J.-P. 2008, *ApJ*, 674, 711
- Sand, D. J., Treu, T., Smith, G. P., & Ellis, R. S. 2004, *ApJ*, 604, 88
- Sérsic, J. L. 1963, *Boletín de la Asociación Argentina de Astronomía La Plata Argentina*, 6, 41
- Skrutskie, M. F., Cutri, R. M., Stiening, R., Weinberg, M. D., Schneider, S., Carpenter, J. M., Beichman, C., Capps, R., Chester, T., Elias, J., Huchra, J., Liebert, J., Lonsdale, C., Monet, D. G., Price, S., Seitzer, P., Jarrett, T., Kirkpatrick, J. D., Gizis, J. E., Howard, E., Evans, T., Fowler, J., Fullmer, L., Hurt, R., Light, R., Kopan, E. L., Marsh, K. A., McCallon, H. L., Tam, R., Van Dyk, S., & Wheelock, S. 2006, *AJ*, 131, 1163
- Sommer-Larsen, J. & Limousin, M. 2009, *arXiv:0906.0573*
- Spekkens, K., Giovanelli, R., & Haynes, M. P. 2005, *AJ*, 129, 2119
- Taylor, J. E. & Navarro, J. F. 2001, *ApJ*, 563, 483
- Trotta, R. 2008, *Contemporary Physics*, 49, 71
- Zait, A., Hoffman, Y., & Shlosman, I. 2008, *ApJ*, 682, 835
- Zappacosta, L., Buote, D. A., Gastaldello, F., Humphrey, P. J., Bullock, J., Brighenti, F., & Mathews, W. 2006, *ApJ*, 650, 777
- Zhao, H. 1996, *MNRAS*, 278, 488